

Quasi-MLE for quadratic ARCH model with long memory

Ieva Grublytė^{1,2}, Donatas Surgailis², Andrius Škarnulis²

September 23, 2015

¹ Université de Cergy Pontoise, Département de Mathématiques, 95302 Cedex, France

² Vilnius University, Institute of Mathematics and Informatics, 08663 Vilnius, Lithuania

Abstract

We discuss parametric quasi-maximum likelihood estimation for quadratic ARCH process with long memory introduced in Doukhan et al. (2015) and Grublytė and Škarnulis (2015) with conditional variance given by a strictly positive quadratic form of observable stationary sequence. We prove consistency and asymptotic normality of the corresponding QMLE estimates, including the estimate of long memory parameter $0 < d < 1/2$. A simulation study of empirical MSE is included.

Keywords: quadratic ARCH process; LARCH model; long memory; parametric estimation; quasi-maximum likelihood

1 Introduction

Recently, Doukhan et al. [4] and Grublytė and Škarnulis [8] discussed a class of quadratic ARCH models of the form

$$r_t = \zeta_t \sigma_t, \quad \sigma_t^2 = \omega^2 + \left(a + \sum_{j=1}^{\infty} b_j r_{t-j}\right)^2 + \gamma \sigma_{t-1}^2, \quad (1.1)$$

where $\gamma, \omega, a, b_j, j \geq 1$ are real parameters. In [8], (1.1) was called the Generalized Quadratic ARCH (GQARCH) model. By iterating the second equation in (1.1), the squared volatility in (1.1) can be written as a quadratic form

$$\sigma_t^2 = \sum_{\ell=0}^{\infty} \gamma^{\ell} \left\{ \omega^2 + \left(a + \sum_{j=1}^{\infty} b_j r_{t-\ell-j}\right)^2 \right\}$$

in lagged variables r_{t-1}, r_{t-2}, \dots , and hence it represents a particular case of Sentana's [13] Quadratic ARCH model with $p = \infty$. The model (1.1) includes the classical Asymmetric GARCH(1,1) process of Engle [5] and the Linear ARCH (LARCH) model of Robinson [11]:

$$r_t = \zeta_t \sigma_t, \quad \sigma_t = a + \sum_{j=1}^{\infty} b_j r_{t-j}. \quad (1.2)$$

Giraitis et al. [7] proved that the squared stationary solution $\{r_t^2\}$ of the LARCH model in (1.2) with b_j decaying as j^{d-1} , $0 < d < 1/2$ may have long memory autocorrelations. For the GQARCH model in (1.1), similar results were established in [4] and [8]. Namely, assume that the parameters $\gamma, \omega, a, b_j, j \geq 1$ in (1.1) satisfy

$$b_j \sim c j^{d-1} \quad (\exists 0 < d < 1/2, c > 0)$$

and

$$B_2^2 \mu_4 K_4 < 1 - \gamma, \quad \gamma \in [0, 1), \quad a \neq 0,$$

where

$$\mu_4 := \mathbb{E} \zeta_0^4, \quad B_2 := \sum_{j=1}^{\infty} b_j^2,$$

and where K_4 is the absolute constant from Rosenthal's inequality in (2.5), below. Then (see [8], Thm.5) there exists a stationary solution of (1.1) with $\mathbb{E} r_t^4 < \infty$ such that

$$\text{cov}(r_0^2, r_t^2) \sim \kappa_1^2 t^{2d-1}, \quad t \rightarrow \infty$$

and

$$n^{-d-1/2} \sum_{t=1}^{[n\tau]} (r_t^2 - \mathbb{E} r_t^2) \rightarrow_{D[0,1]} \kappa_2 W_{d+(1/2)}(\tau), \quad n \rightarrow \infty,$$

where $W_{d+(1/2)}$ is a fractional Brownian motion with Hurst parameter $H = d + (1/2) \in (1/2, 1)$ and $\kappa_i > 0, i = 1, 2$ are some constants; $\rightarrow_{D[0,1]}$ stands for the weak convergence in the Skorohod space $D[0, 1]$.

As noted in [4], [8], the GQARCH model of (1.1) and the LARCH model of (1.2) have similar long memory and leverage properties and both can be used for modelling of financial data with the above properties. The main disadvantage of the latter model vs. the former one seems to be the fact that volatility σ_t in (1.2) may assume negative values and is not separated from below by positive constant $c > 0$ as in the case of (1.1). The standard quasi-maximum likelihood (QML) approach to estimation of LARCH parameters is inconsistent and other estimation methods were developed in Beran and Schützner [1], Francq and Zakoian [6], Levine et al. [9], Truquet [14].

The present paper discusses QML estimation for the 5-parametric GQARCH model

$$\sigma_t^2(\theta) = \sum_{\ell=0}^{\infty} \gamma^\ell \{ \omega^2 + (a + c \sum_{j=1}^{\infty} j^{d-1} r_{t-\ell-j})^2 \}, \quad (1.3)$$

depending on parameter $\theta = (\gamma, \omega, a, d, c), 0 < \gamma < 1, \omega > 0, a \neq 0, c \neq 0$ and $d \in (0, 1/2)$. The parametric form $b_j = c j^{d-1}$ of moving-average coefficients in (1.3) is the same as in Beran and Schützner [1] for the LARCH model. Similarly as in [1] we discuss the QML estimator $\hat{\theta}_n := \arg \min_{\theta \in \Theta} L_n(\theta)$, $L_n(\theta) := \frac{1}{n} \sum_{t=1}^n \left(\frac{r_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta) \right)$ involving exact conditional variance in (1.3) depending on infinite past $r_s, -\infty < s < t$, and its more realistic version $\tilde{\theta}_n := \arg \min_{\theta \in \Theta} \tilde{L}_n(\theta)$, obtained by replacing the $\sigma_t^2(\theta)$'s in (1.3) by $\tilde{\sigma}_t^2(\theta)$ depending only $r_s, 1 \leq s < t$ (see Sec. 3 for the definition). It should be noted that the

QML function in [1] is modified to avoid the degeneracy of σ_t^{-1} in (1.2), by introducing an additional tuning parameter $\epsilon > 0$ which affects the performance of the estimator and whose choice is a non-trivial task. For the GQARCH model (1.3) with $\omega > 0$ the above degeneracy problem does not occur and we deal with unmodified QMLE in contrast to [1]. We also note that our proofs use different techniques from [1]. Particularly, the method of orthogonal Volterra expansions of the LARCH model used in [1] is not applicable for model (1.3); see ([4], Example 1).

This paper is organized as follows. Sec. 2 presents some results of [8] about the existence and properties of stationary solution of the GQARCH equations in (1.1). In Sec. 3 we define several QMLE estimators of parameter θ in (1.3). Sec. 4 presents the main results of the paper devoted to consistency and asymptotic normality of the QML estimators. Finite sample performance of these estimators is investigated in the simulation study of Sec. 5. All proofs are relegated to Sec. 6.

2 Properties of stationary solution

In this sec. we recall some facts from [8] about stationary solution of (1.1). First, we give the definition of it. Let $\mathcal{F}_t = \sigma(\zeta_s, s \leq t), t \in \mathbb{Z}$ be the sigma-field generated by $\zeta_s, s \leq t$.

Definition 2.1 *By stationary solution of (1.1) we mean a stationary and ergodic martingale difference sequence $\{r_t, \mathcal{F}_t, t \in \mathbb{Z}\}$ with $E r_t^2 < \infty, E[r_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2$ such that for any $t \in \mathbb{Z}$ the series $X_t := \sum_{s < t} b_{t-s} r_s$ converges in L^2 , the series $\sigma_t^2 = \sum_{\ell=0}^{\infty} \gamma^\ell (\omega^2 + (a + X_{t-\ell})^2)$ converges in L^1 and (1.1) holds.*

For real $p \geq 2$, define

$$B_p := \left(\sum_{j=1}^{\infty} b_j^2 \right)^{p/2}, \quad B_{p,\gamma} := B_p / (1 - \gamma). \quad (2.4)$$

We use the following moment inequality.

Proposition 2.2 *Let $p \geq 2$ and $\{Y_j\}$ be a martingale difference sequence such that $E|Y_j|^p < \infty; E[Y_j | Y_1, \dots, Y_{j-1}] = 0, j = 2, 3, \dots$. Then there exists a constant K_p depending only on p and such that*

$$E \left| \sum_{j=1}^{\infty} Y_j \right|^p \leq K_p \left(\sum_{j=1}^{\infty} (E|Y_j|^p)^{2/p} \right)^{p/2}. \quad (2.5)$$

Remark 2.3 Inequality (2.5) is trivial for $p = 2, K_2 = 1$. For $p > 2$, (2.5) is a consequence of the Burkholder and Rosenthal inequality (see [3], [12]). Osękowski [10] proved that $K_p^{1/p} \leq 4(\frac{p}{4} + 1)^{1/p} (1 + \frac{p}{\log(p/2)})$, in particular, $K_4^{1/4} \leq 32.207$.

Proposition 2.4 ([8]) *Let $\gamma \in [0, 1)$ and $\{\zeta_t\}$ be an i.i.d. sequence with zero mean and $|\mu|_p := E|\zeta_0|^p < \infty$ for some $p \geq 2$. Assume that*

$$K_p |\mu|_p B_{p,\gamma} < 1, \quad (2.6)$$

where $B_{p,\gamma}$ is defined in (2.4) and K_p is the absolute constant in (2.5). Then there exists a unique stationary solution $\{r_t\}$ of (1.1) such that the series $X_t = \sum_{j=1}^{\infty} b_j r_{t-j}$ converges in L^p and

$$\mathbb{E}|r_t|^p \leq C(1 + \mathbb{E}|X_t|^p) \quad \text{and} \quad \mathbb{E}|X_t|^p \leq \frac{CB_p}{1 - K_p|\mu|_p B_{p,\gamma}},$$

where $C > 0$ is a constant independent of $\{b_j\}, p, \gamma$, and the distribution of ζ_0 . Moreover, for $p = 2$ condition (2.6), or

$$B_2 = \sum_{j=1}^{\infty} b_j^2 < 1 - \gamma \tag{2.7}$$

is necessary for the existence of a stationary L^2 -solution of (1.1).

3 QML estimators

The following assumptions on the parametric GQARCH model in (1.3) are imposed.

Assumption (A) $\{\zeta_t\}$ is a standardized i.i.d. sequence with $\mathbb{E}\zeta_t = 0, \mathbb{E}\zeta_t^2 = 1$.

Assumption (B) $\Theta \subset \mathbb{R}^5$ is a compact set of parameters $\theta = (\gamma, \omega, a, d, c)$ defined by

- (i) $\gamma \in [\gamma_1, \gamma_2]$ with $0 < \gamma_1 < \gamma_2 < 1$;
- (ii) $\omega \in [\omega_1, \omega_2]$ with $0 < \omega_1 < \omega_2 < \infty$;
- (iii) $a \in [a_1, a_2]$ with $-\infty < a_1 < a_2 < \infty$;
- (iv) $d \in [d_1, d_2]$ with $0 < d_1 < d_2 < 1/2$;
- (v) $c \in [c_1, c_2]$ with $0 < c_i = c_i(d, \gamma) < \infty$, $c_1 < c_2$ such that $B_2 = c^2 \sum_{j=1}^{\infty} j^{2(d-1)} < 1 - \gamma$ for any $c \in [c_1, c_2], \gamma \in [\gamma_1, \gamma_2], d \in [d_1, d_2]$.

We assume that the observations $\{r_t, 1 \leq t \leq n\}$ follow the model in (1.1) with the true parameter $\theta_0 = (\gamma_0, \omega_0, a_0, d_0, c_0)$ belonging to the interior Θ_0 of Θ in Assumption (B). The restriction on parameter c in (v) is due to condition (2.7). The QML estimator of $\theta \in \Theta$ is defined as

$$\hat{\theta}_n := \arg \min_{\theta \in \Theta} L_n(\theta) \tag{3.8}$$

where

$$L_n(\theta) = \frac{1}{n} \sum_{t=1}^n \left(\frac{r_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta) \right), \tag{3.9}$$

and $\sigma_t^2(\theta)$ is defined in (1.3), viz.,

$$\begin{aligned} \sigma_t^2(\theta) &= \sum_{\ell=0}^{\infty} \gamma^\ell \{ \omega^2 + (a + cY_{t-\ell}(d))^2 \}, \quad \text{where} \\ Y_t(d) &:= \sum_{j=1}^{\infty} j^{d-1} r_{t-j}. \end{aligned} \tag{3.10}$$

Note the definitions in (3.8)-(3.10) depend on (unobserved) $r_s, s \leq 0$ and therefore the estimator in (3.8) is usually referred to as the QMLE given infinite past [1]. A more realistic version of (3.8) is defined as

$$\tilde{\theta}_n := \arg \min_{\theta \in \Theta} \tilde{L}_n(\theta), \quad (3.11)$$

where

$$\begin{aligned} \tilde{L}_n(\theta) &:= \frac{1}{n} \sum_{t=1}^n \left(\frac{r_t^2}{\tilde{\sigma}_t^2(\theta)} + \log \tilde{\sigma}_t^2(\theta) \right), \quad \text{where} \\ \tilde{\sigma}_t^2(\theta) &:= \sum_{\ell=0}^{t-1} \gamma^\ell \{ \omega^2 + (a + c\tilde{Y}_{t-\ell}(d))^2 \}, \quad \tilde{Y}_t(d) := \sum_{j=1}^{t-1} j^{d-1} r_{t-j}. \end{aligned} \quad (3.12)$$

Note all quantities in (3.12) depend only on $r_s, 1 \leq t \leq n$, hence (3.11) is called the QMLE given finite past. The QML functions in (3.9) and (3.12) can be written as

$$L_n(\theta) = \frac{1}{n} \sum_{t=1}^n l_t(\theta) \quad \text{and} \quad \tilde{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{l}_t(\theta),$$

respectively, where

$$l_t(\theta) := \frac{r_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta), \quad \tilde{l}_t(\theta) := \frac{r_t^2}{\tilde{\sigma}_t^2(\theta)} + \log \tilde{\sigma}_t^2(\theta). \quad (3.13)$$

Finally, following [1] we define a truncated version of (3.11) involving the last $O(n^\beta)$ quasi-likelihoods $\tilde{l}_t(\theta), n - [n^\beta] < t \leq n$, as follows:

$$\tilde{\theta}_n^{(\beta)} := \arg \min_{\theta \in \Theta} \tilde{L}_n^{(\beta)}(\theta), \quad \tilde{L}_n^{(\beta)}(\theta) := \frac{1}{[n^\beta]} \sum_{t=n-[n^\beta]+1}^n \tilde{l}_t(\theta). \quad (3.14)$$

where $0 < \beta < 1$ is a ‘bandwidth parameter’. Note that for any $t \in \mathbb{Z}$ and $\theta_0 = (\gamma_0, \omega_0, a_0, d_0, c_0) \in \Theta$, the random functions $Y_t(d)$ and $\tilde{Y}_t(d)$ in (3.10) and (3.12) are infinitely differentiable w.r.t. $d \in (0, 1/2)$ a.s. Hence using the explicit form of $\sigma_t^2(\theta)$ and $\tilde{\sigma}_t^2(\theta)$, it follows that $\sigma_t^2(\theta), \tilde{\sigma}_t^2(\theta), l_t(\theta), \tilde{l}_t(\theta), L_n(\theta), \tilde{L}_n(\theta), \tilde{L}_n^{(\beta)}(\theta)$ etc. are all infinitely differentiable w.r.t. $\theta \in \Theta_0$ a.s. We use the notation

$$L(\theta) := \mathbb{E} L_n(\theta) = \mathbb{E} l_t(\theta) \quad (3.15)$$

and

$$A(\theta) := \mathbb{E} [\nabla^T l_t(\theta) \nabla l_t(\theta)] \quad \text{and} \quad B(\theta) := \mathbb{E} [\nabla^T \nabla l_t(\theta)], \quad (3.16)$$

where $\nabla = (\partial/\partial\theta_1, \dots, \partial/\partial\theta_5)$ and the superscript T stands for transposed vector. Particularly, $A(\theta)$ and $B(\theta)$ are 5×5 -matrices. By Lemma 4.1, the expectations in (3.16) are well-defined for any $\theta \in \Theta$ under condition $\mathbb{E} r_0^4 < \infty$. We have

$$B(\theta) = \mathbb{E} [\sigma_t^{-4}(\theta) \nabla^T \sigma_t^2(\theta) \nabla \sigma_t^2(\theta)] \quad \text{and} \quad A(\theta) = \kappa_4 B(\theta) \quad (3.17)$$

where $\kappa_4 := \mathbb{E}(\zeta_0^2 - 1)^2 > 0$.

4 Main results

Everywhere in this section $\{r_t\}$ is a stationary solution of model (1.3) as defined in Definition 2.1 and satisfying Assumptions (A) and (B) of the previous section. As usual, all expectations are taken with respect to the true value $\theta_0 = (\gamma_0, \omega_0, a_0, d_0, c_0) \in \Theta_0$, where Θ_0 is the interior of the parameter set $\Theta \subset \mathbb{R}^5$.

Theorem 4.1 (i) Let $E|r_t|^3 < \infty$. Then $\hat{\theta}_n$ in (3.8) is a strongly consistent estimator of θ_0 , i.e.

$$\hat{\theta}_n \xrightarrow{a.s.} \theta_0.$$

(ii) Let $E|r_t|^5 < \infty$. Then $\hat{\theta}_n$ in (3.8) is asymptotically normal:

$$n^{1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{law} N(0, \Sigma(\theta_0)), \quad (4.18)$$

where $\Sigma(\theta_0) := B^{-1}(\theta_0)A(\theta_0)B^{-1}(\theta_0) = \kappa_4 B^{-1}(\theta_0)$ and matrices $A(\theta), B(\theta)$ are defined in (3.17).

The following theorem gives asymptotic properties of ‘finite past’ estimators $\tilde{\theta}_n$ and $\tilde{\theta}_n^{(\beta)}$ defined in (3.11) and (3.14), respectively.

Theorem 4.2 (i) Let $E|r_t|^3 < \infty$ and $0 < \beta < 1$. Then

$$E|\tilde{\theta}_n - \theta_0| \rightarrow 0 \quad \text{and} \quad E|\tilde{\theta}_n^{(\beta)} - \theta_0| \rightarrow 0.$$

(ii) Let $E|r_t|^5 < \infty$ and $0 < \beta < 1 - 2d_0$. Then

$$n^{\beta/2}(\tilde{\theta}_n^{(\beta)} - \theta_0) \xrightarrow{law} N(0, \Sigma(\theta_0)), \quad (4.19)$$

where $\Sigma(\theta_0)$ is the same as in Theorem 4.1.

The asymptotic results in Theorems 4.1 and 4.2 are similar to the results of ([1], Thm. 1-4) pertaining to the 3-parametric LARCH model in (1.2) with $b_j = cj^{d-1}$, except that [1] deal with a modified QMLE involving a ‘tuning parameter’ $\epsilon > 0$. Theorems 4.1 and 4.2 are based on subsequent Lemmas 4.1-4.4 which describe properties of the likelihood processes defined in (3.9), (3.12) and (3.13). As noted in Sec. 1, our proofs use different techniques from [1] which rely on explicit Volterra series representation of stationary solution of the LARCH model.

For multi-index $\mathbf{i} = (i_1, \dots, i_5) \in \mathbb{N}^5$, $\mathbf{i} \neq \mathbf{0} = (0, \dots, 0)$, $|\mathbf{i}| := i_1 + \dots + i_5$, denote partial derivative $\partial^{\mathbf{i}} := \partial^{|\mathbf{i}|} / \prod_{j=1}^5 \partial^{i_j} \theta_{i_j}$.

Lemma 4.1 Let $E|r_t|^{2+p} < \infty$, for some integer $p \geq 1$. Then for any $\mathbf{i} \in \mathbb{N}^5$, $0 < |\mathbf{i}| \leq p$,

$$E \sup_{\theta \in \Theta} |\partial^{\mathbf{i}} l_t(\theta)| < \infty. \quad (4.20)$$

Moreover, if $E|r_t|^{2+p+\epsilon} < \infty$ for some $\epsilon > 0$ and $p \in \mathbb{N}$ then for any $\mathbf{i} \in \mathbb{N}^5$, $0 \leq |\mathbf{i}| \leq p$

$$E \sup_{\theta \in \Theta} |\partial^{\mathbf{i}} (l_t(\theta) - \tilde{l}_t(\theta))| \rightarrow 0, \quad t \rightarrow \infty. \quad (4.21)$$

Lemma 4.2 *The function $L(\theta), \theta \in \Theta$ in (3.15) is bounded and continuous. Moreover, it attains its unique minimum at $\theta = \theta_0$.*

Lemma 4.3 *Let $Er_0^4 < \infty$. Then matrices $A(\theta)$ and $B(\theta)$ in (3.16) are well-defined and strictly positive definite for any $\theta \in \Theta$.*

Write $|\cdot|$ for the Euclidean norm in \mathbb{R}^5 and in $\mathbb{R}^5 \otimes \mathbb{R}^5$ (the matrix norm).

Lemma 4.4 (i) *Let $E|r_t|^3 < \infty$. Then*

$$\sup_{\theta \in \Theta} |L_n(\theta) - L(\theta)| \xrightarrow{a.s.} 0 \quad \text{and} \quad E \sup_{\theta \in \Theta} |L_n(\theta) - \tilde{L}_n(\theta)| \rightarrow 0. \quad (4.22)$$

(ii) *Let $Er_t^4 < \infty$. Then $\nabla L(\theta) = E \nabla l_t(\theta)$ and*

$$\sup_{\theta \in \Theta} |\nabla L_n(\theta) - \nabla L(\theta)| \xrightarrow{a.s.} 0 \quad \text{and} \quad E \sup_{\theta \in \Theta} |\nabla L_n(\theta) - \nabla \tilde{L}_n(\theta)| \rightarrow 0. \quad (4.23)$$

(iii) *Let $E|r_t|^5 < \infty$. Then $\nabla^T \nabla L(\theta) = E \nabla^T \nabla l_t(\theta) = B(\theta)$ (see (3.16)) and*

$$\sup_{\theta \in \Theta} |\nabla^T \nabla L_n(\theta) - \nabla^T \nabla L(\theta)| \xrightarrow{a.s.} 0, \quad (4.24)$$

$$E \sup_{\theta \in \Theta} |\nabla^T \nabla L_n(\theta) - \nabla^T \nabla \tilde{L}_n(\theta)| \rightarrow 0. \quad (4.25)$$

5 Simulation study

In this section we present a short simulation study of the performance of the QMLE for the GQARCH model in (1.3). The GQARCH model in (1.3) was simulated with i.i.d. standard normal innovations $\{\zeta_t\}$. The QMLE procedure was evaluated for medium-term ($n = 1000$) and long-term ($n = 5000$) samples. The process was generated for $-n \leq t \leq n$ using the recurrent formula in (1.1) with appropriately truncated sum $\sum_{j=1}^{\min(n, t+n)}$ and zero initial condition $\sigma_{-n-1} = 0$. The QMLE estimation used generated time series $r_t, 1 \leq t \leq n$ with $r_t, -n \leq t \leq 0$ as the pre-sample. The numerical optimization procedure minimized the QML function:

$$L_n = \frac{1}{n} \sum_{t=1}^n \left(\frac{r_t^2}{\sigma_t^2} + \log \sigma_t^2 \right),$$

with

$$r_t = \zeta_t \sigma_t, \quad \sigma_t^2 = \omega^2 + \left(a + c \sum_{j=1}^n j^{d-1} r_{t-j} \right)^2 + \gamma \sigma_{t-1}^2, \quad t = 1, \dots, n.$$

Finite-sample performance of the QML estimator is studied for fixed values of parameters $\gamma_0 = 0.7, a_0 = -0.2, c_0 = 0.2$ and different values of $\omega_0 = 0.1, 0.01, 0.001$ and the long memory parameter $d_0 = 0.1, 0.2, 0.3, 0.4$. The above choice of $\theta_0 = (\gamma_0, \omega_0, a_0, c_0, d_0)$ can be explained by the observation that the QML estimation of γ_0, a_0, c_0 appears to be more accurate and stable in comparison with estimation of ω_0 and d_0 . The very small values of ω_0 in our experiment reflect the fact that in most real data studied by us, the estimated

QML value of ω_0 was less than 0.05. The presence of $\omega_0 > 0$ in the GQARCH model in (1.3) is very important for consistency of the QMLE procedure, by guaranteeing that $\sigma_t^2(\theta)$ is separated from zero. A similar role is played by the ‘tuning parameter’ $\epsilon > 0$ in the LARCH estimation in [1], except that $\omega_0 > 0$ is estimated in (1.3) and not *ad hoc* imposed as $\epsilon > 0$ in [1].

The numerical QML minimization was performed using the MATLAB language for technical computing, under the following constraints:

$$\begin{aligned} 0.001 \leq \gamma \leq 0.9, \quad 0 \leq \omega \leq 2, \quad -2 \leq a \leq 2, \quad 0 \leq d \leq 0.5, \\ (0.05 - \gamma) \vee (\gamma/999) \leq c^2 \zeta(2(1 - d)) \leq (0.99 - \gamma) \wedge (99\gamma), \end{aligned} \quad (5.26)$$

where $\zeta(z) = \sum_{j=1}^{\infty} j^{-z}$ is the Riemann zeta function. The last constraint in (5.26) guarantees Assumption (B) (v) with appropriate $0 < c_i(d, \gamma), i = 1, 2$.

The results of the simulation experiment are presented in Table 1, which shows the sample R(oot)MSEs of the QML estimates $\hat{\theta}_n = (\hat{\gamma}_n, \hat{\omega}_n, \hat{a}_n, \hat{c}_n, \hat{d}_n)$ with 100 independent replications, for two sample lengths $n = 1000$ and $n = 5000$ and the above choices of $\theta_0 = (\gamma_0, \omega_0, a_0, c_0, d_0)$. Our observations from Table 1 are summarized below.

1. All RMSEs decrease as n increases. The convergence rate of estimates seems quite good overall.
2. Parameter γ_0 is estimated rather accurately. E.g., for $n = 5000$ $\text{RMSE}(\hat{\gamma}_n)$ is very stable for all values of ω_0 and d_0 .
3. The previous conclusion generally applies also to the QML estimates \hat{a}_n, \hat{c}_n and \hat{d}_n except that their RMSE markedly increases when $d_0 = 0.4$.
4. The QML estimate of $\omega_0 \leq 0.01$ seems to have a ‘constant’ bias $\approx 0.02 \div 0.03$ for all values of d_0 with $n = 5000$.

6 Proofs

Proof of Lemma 4.1. We use the following (Faà di Bruno) differentiation rule:

$$\begin{aligned} \partial^{\mathbf{i}} \sigma_t^{-2}(\theta) &= \sum_{\nu=1}^{|\mathbf{i}|} (-1)^{\nu} \nu! \sigma_t^{-2(1+\nu)}(\theta) \sum_{\mathbf{j}_1 + \dots + \mathbf{j}_{\nu} = \mathbf{i}} \chi_{\mathbf{j}_1, \dots, \mathbf{j}_{\nu}} \prod_{k=1}^{\nu} \partial^{\mathbf{j}_k} \sigma_t^2(\theta), \\ \partial^{\mathbf{i}} \log \sigma_t^2(\theta) &= \sum_{\nu=1}^{|\mathbf{i}|} (-1)^{\nu-1} (\nu-1)! \sigma_t^{-2\nu}(\theta) \sum_{\mathbf{j}_1 + \dots + \mathbf{j}_{\nu} = \mathbf{i}} \chi_{\mathbf{j}_1, \dots, \mathbf{j}_{\nu}} \prod_{k=1}^{\nu} \partial^{\mathbf{j}_k} \sigma_t^2(\theta), \end{aligned} \quad (6.27)$$

where the sum $\sum_{\mathbf{j}_1 + \dots + \mathbf{j}_{\nu} = \mathbf{i}}$ is taken over decompositions of \mathbf{i} into a sum of ν multi-indices $\mathbf{j}_k \neq \mathbf{0}, k = 1, \dots, \nu$, and $\chi_{\mathbf{j}_1, \dots, \mathbf{j}_{\nu}}$ is a combinatorial factor depending only on $\mathbf{j}_k, 1 \leq k \leq \nu$.

Let us prove (4.20). We have $|\partial^{\mathbf{i}} l_t(\theta)| \leq r_t^2 |\partial^{\mathbf{i}} \sigma_t^{-2}(\theta)| + |\partial^{\mathbf{i}} \log \sigma_t^2(\theta)|$. Hence using (6.27) and the fact that $\sigma_t^2(\theta) \geq \omega^2/(1-\gamma) \geq \omega_1^2/(1-\gamma_2) > 0$ we obtain

$$\sup_{\theta \in \Theta} |\partial^{\mathbf{i}} l_t(\theta)| \leq C(r_t^2 + 1) \sum_{\nu=1}^{|\mathbf{i}|} \sum_{\mathbf{j}_1 + \dots + \mathbf{j}_\nu = \mathbf{i}} \prod_{k=1}^{\nu} \sup_{\theta \in \Theta} (|\partial^{\mathbf{j}_k} \sigma_t^2(\theta)| / \sigma_t(\theta)).$$

Therefore by Hölder's inequality

$$\begin{aligned} \mathbb{E} \sup_{\theta \in \Theta} |\partial^{\mathbf{i}} l_t(\theta)| &\leq C(\mathbb{E}(r_t^2 + 1)^{(2+p)/2})^{2/(2+p)} \\ &\times \sum_{\nu=1}^{|\mathbf{i}|} \sum_{\mathbf{j}_1 + \dots + \mathbf{j}_\nu = \mathbf{i}} \prod_{k=1}^{\nu} \mathbb{E}^{1/q_k} \left(\sup_{\theta \in \Theta} |\partial^{\mathbf{j}_k} \sigma_t^2(\theta)| / \sigma_t(\theta) \right)^{q_k}, \end{aligned} \quad (6.28)$$

where $\sum_{j=1}^{\nu} 1/q_j \leq p/(2+p)$. Note $|\mathbf{i}| = \sum_{k=1}^{\nu} |\mathbf{j}_k|$ and hence the choice $q_k = (2+p)/|\mathbf{j}_k|$ satisfies $\sum_{j=1}^{\nu} 1/q_j = \sum_{k=1}^{\nu} |\mathbf{j}_k|/(2+p) \leq p/(2+p)$. Using (6.28) and condition $\mathbb{E}|r_t|^{2+p} \leq C$, relation (4.20) follows from

$$\mathbb{E} \sup_{\theta \in \Theta} (|\partial^{\mathbf{j}} \sigma_t^2(\theta)| / \sigma_t(\theta))^{(2+p)/|\mathbf{j}|} < \infty \quad (6.29)$$

for any multi-index $\mathbf{j} \in \mathbb{N}^5$, $1 \leq |\mathbf{j}| \leq p$.

Consider first the case $|\mathbf{j}| = 1$, or the partial derivative $\partial_i \sigma_t^2(\theta) = \partial \sigma_t^2(\theta) / \partial \theta_i$, $1 \leq i \leq 5$. We have

$$\partial_i \sigma_t^2(\theta) = \begin{cases} \sum_{\ell=1}^{\infty} \ell \gamma^{\ell-1} \{\omega^2 + (a + cY_{t-\ell}(d))^2\}, & \theta_i = \gamma, \\ \sum_{\ell=0}^{\infty} \gamma^{\ell} 2\omega, & \theta_i = \omega, \\ \sum_{\ell=0}^{\infty} \gamma^{\ell} 2(a + cY_{t-\ell}(d)), & \theta_i = a, \\ \sum_{\ell=0}^{\infty} \gamma^{\ell} 2(a + cY_{t-\ell}(d))Y_{t-\ell}(d), & \theta_i = c, \\ \sum_{\ell=0}^{\infty} \gamma^{\ell} 2c(a + cY_{t-\ell}(d))\partial_d Y_{t-\ell}(d), & \theta_i = d. \end{cases} \quad (6.30)$$

We claim that there exist $C > 0, 0 < \bar{\gamma} < 1$ such that

$$\begin{aligned} \sup_{\theta \in \Theta} \left| \frac{\partial_i \sigma_t^2(\theta)}{\sigma_t(\theta)} \right| &\leq C(1 + J_{t,0} + J_{t,1}), \quad i = 1, \dots, 5, \quad \text{where} \\ J_{t,0} &:= \sum_{\ell=0}^{\infty} \bar{\gamma}^{\ell} \sup_{d \in [d_1, d_2]} |Y_{t-\ell}(d)|, \quad J_{t,1} := \sum_{\ell=0}^{\infty} \bar{\gamma}^{\ell} \sup_{d \in [d_1, d_2]} |\partial_d Y_{t-\ell}(d)|. \end{aligned} \quad (6.31)$$

Consider (6.31) for $\theta_i = \gamma$. Using $\ell^2 \gamma^{\ell-2} \leq C \bar{\gamma}^{\ell}$ for all $\ell \geq 1, \gamma \in [\gamma_1, \gamma_2] \subset (0, 1)$ and some $C > 0, 0 < \bar{\gamma} < 1$ together with Assumption (B) and Cauchy inequality, we obtain $|\partial_{\gamma} \sigma_t^2(\theta)| / \sigma_t(\theta) \leq (\sum_{\ell=1}^{\infty} \ell^2 \gamma^{\ell-2} \{\omega^2 + (a + cY_{t-\ell}(d))^2\})^{1/2} \leq C(1 + J_{t,0})$ uniformly in $\theta \in \Theta$, proving (6.31) for $\theta_i = \gamma$. Similarly, $|\partial_c \sigma_t^2(\theta)| / \sigma_t(\theta) \leq C(1 + J_{t,0})$ and $|\partial_d \sigma_t^2(\theta)| / \sigma_t(\theta) \leq C(1 + J_{t,1})$. Finally, for $\theta_i = \omega$ and $\theta_i = a$, (6.31) is immediate from (6.30), proving (6.31).

With (6.31) in mind, (6.29) for $|\mathbf{j}| = 1$ follows from

$$\mathbb{E} J_{t,i}^{2+p} = \mathbb{E} \left(\sum_{\ell=0}^{\infty} \bar{\gamma}^{\ell} \sup_{d \in [d_1, d_2]} |\partial_d^i Y_{t-\ell}(d)| \right)^{2+p} < \infty, \quad i = 0, 1. \quad (6.32)$$

Using Minkowski's inequality and stationarity of $\{Y_t(d)\}$ we obtain $E^{1/(2+p)} J_{t,i}^{2+p} \leq \sum_{\ell=0}^{\infty} \bar{\gamma}^{\ell} E^{1/(2+p)} \sup_d |\partial_d^i Y_{t-\ell}(d)|^{2+p} \leq C(E \sup_d |\partial_d^i Y_t(d)|^{2+p})^{1/(2+p)}$, where $\partial_d^i Y_t(d) = \sum_{j=1}^{\infty} \partial_d^i j^{d-1} r_{t-j}$. Hence using ([1], Lemma 1 (b)) and the inequality $xy \leq x^q/q + y^{q'}/q'$, $x, y > 0$, $1/q + 1/q' = 1$ we obtain

$$\begin{aligned} \sum_{i=0}^1 E J_{t,i}^{2+p} &\leq C \sum_{i=0}^1 E \sup_{d \in [d_1, d_2]} |\partial_d^i Y_t(d)|^{2+p} \\ &\leq C \sum_{i=0}^2 \sup_{d \in [d_1, d_2]} E |\partial_d^i Y_t(d)|^{2+p} < \infty \end{aligned} \quad (6.33)$$

since $\sup_{d \in [d_1, d_2]} E |\partial_d^i Y_t(d)|^{2+p} \leq C \sup_{d \in [d_1, d_2]} \left(\sum_{j=1}^{\infty} (\partial_d^i j^{d-1})^2 (E |r_{t-j}|^{2+p})^{2/(2+p)} \right)^{(2+p)/2} < \infty$ according to condition $E |r_t|^{2+p} < C$, Rosenthal's inequality in (2.5) and the fact that $\sup_{d \in [d_1, d_2]} \sum_{j=1}^{\infty} (\partial_d^i j^{d-1})^2 \leq \sup_{d \in [d_1, d_2]} \sum_{j=1}^{\infty} j^{2(d-1)} (1 + \log^2 j)^2 < C$, $i = 0, 1, 2$. This proves (6.29) for $|j| = 1$.

The proof of (6.29) for $2 \leq |j| \leq p$ is simpler since it reduces to

$$E \sup_{\theta \in \Theta} |\partial^{\dot{j}} \sigma_t^2(\theta)|^{(p+2)/2} < \infty, \quad 2 \leq |j| \leq p. \quad (6.34)$$

Recall $\theta_1 = \gamma$ and $j' := j - (j_1, 0, 0, 0, 0) = (0, j_2, j_3, j_4, j_5)$. If $j' = \mathbf{0}$ then $\sup_{\theta \in \Theta} |\partial^{\dot{j}} \sigma_t^2(\theta)| \leq C J_{t,0}$ follows as in (6.31) implying (6.34) as in (6.33) above. Next, let $j' \neq \mathbf{0}$. Denote

$$Q_t^2(\theta) := \omega^2 + (a + cY_t(d))^2 \quad (6.35)$$

so that $\sigma_t^2(\theta) = \sum_{\ell=0}^{\infty} \gamma^{\ell} Q_{t-\ell}^2(\theta)$. We have with $m := j_1 \geq 0$ that $|\partial^{\dot{j}} \sigma_t^2(\theta)| \leq \sum_{\ell=m}^{\infty} (\ell! / (\ell - m)!) \gamma^{\ell-m} |\partial^{\dot{j}'} Q_{t-\ell}^2(\theta)|$ and (6.29) follows from

$$E \sup_{\theta \in \Theta} |\partial^{\dot{j}} Q_t^2(\theta)|^{(p+2)/2} < \infty. \quad (6.36)$$

For $j_2 \neq 0$ (recall $\theta_2 = \omega$) the derivative in (6.36) is trivial so that it suffices to check (6.36) for $j_1 = 0$ only. Then applying Faà di Bruno's rule we get

$$|\partial^{\dot{j}} Q_t^2(\theta)|^{(p+2)/2} \leq C \sum_{\dot{j}_1 + \dot{j}_2 = \dot{j}} |\partial^{\dot{j}_1} (a + cY_t(d))|^{(p+2)/2} |\partial^{\dot{j}_2} (a + cY_t(d))|^{(p+2)/2}$$

and hence (6.36) reduces to

$$E \sup_{\theta \in \Theta} |\partial^{\dot{j}} (a + cY_t(d))|^{p+2} < \infty, \quad 0 \leq |j| \leq p,$$

whose proof is similar to (6.32) above. This ends the proof of (4.20).

The proof of (4.21) is similar. We have $|\partial^{\dot{i}}(l_t(\theta) - \tilde{l}_t(\theta))| \leq r_t^2 |\partial^{\dot{i}}(\sigma_t^{-2}(\theta) - \tilde{\sigma}_t^{-2}(\theta))| + |\partial^{\dot{i}}(\log \sigma_t^2(\theta) - \log \tilde{\sigma}_t^2(\theta))|$. Hence, using Hölder's inequality similarly as in the proof (4.20) it suffices to show

$$E \sup_{\theta \in \Theta} |\partial^{\dot{i}}(\sigma_t^{-2}(\theta) - \tilde{\sigma}_t^{-2}(\theta))|^{\frac{p+2}{p}} \rightarrow 0 \quad \text{and} \quad E \sup_{\theta \in \Theta} |\partial^{\dot{i}}(\log \sigma_t^2(\theta) - \log \tilde{\sigma}_t^2(\theta))|^{\frac{p+2}{p}} \rightarrow 0. \quad (6.37)$$

Below, we prove the first relation in (6.37) only, the proof of the second one being analogous.

Using the differentiation rule in (6.27) we have that

$$|\partial^{\mathbf{j}}(\sigma_t^{-2}(\theta) - \tilde{\sigma}_t^{-2}(\theta))| \leq C \sum_{\nu=1}^{|\mathbf{j}|} \sum_{\mathbf{j}_1+\dots+\mathbf{j}_\nu=\mathbf{j}} |W_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta) - \widetilde{W}_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta)|,$$

where

$$\begin{aligned} W_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta) &:= \sigma_t^{-2(1+\nu)}(\theta) \prod_{k=1}^{\nu} \partial^{\mathbf{j}_k} \sigma_t^2(\theta), \\ \widetilde{W}_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta) &:= \tilde{\sigma}_t^{-2(1+\nu)}(\theta) \prod_{k=1}^{\nu} \partial^{\mathbf{j}_k} \tilde{\sigma}_t^2(\theta). \end{aligned}$$

Whence, (6.37) follows from

$$\sup_{\theta \in \Theta} |W_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta) - \widetilde{W}_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta)| \xrightarrow{p} 0, \quad t \rightarrow \infty \quad (6.38)$$

and

$$\mathbb{E} \sup_{\theta \in \Theta} (|W_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta)| + |\widetilde{W}_t^{\mathbf{j}_1,\dots,\mathbf{j}_\nu}(\theta)|)^{(p+2+\epsilon)/p} \leq C < \infty \quad (6.39)$$

for some constants $\epsilon > 0$ and $C > 0$ independent of t . In turn, (6.38) and (6.39) follow from

$$\sup_{\theta \in \Theta} |\partial^{\mathbf{j}}(\sigma_t^2(\theta) - \tilde{\sigma}_t^2(\theta))| \xrightarrow{p} 0, \quad t \rightarrow \infty \quad (6.40)$$

and

$$\begin{aligned} \mathbb{E} \sup_{\theta \in \Theta} (|\partial^{\mathbf{j}} \sigma_t^2(\theta)| / \sigma_t(\theta))^{(2+p+\epsilon)/|\mathbf{j}|} &< C, \\ \mathbb{E} \sup_{\theta \in \Theta} (|\partial^{\mathbf{j}} \tilde{\sigma}_t^2(\theta)| / \tilde{\sigma}_t(\theta))^{(2+p+\epsilon)/|\mathbf{j}|} &< C, \end{aligned} \quad (6.41)$$

for any multi-index \mathbf{j} such that $|\mathbf{j}| \geq 0$ and $1 \leq |\mathbf{j}| \leq p$, respectively.

Using condition $\mathbb{E}|r_t|^{2+p+\epsilon} < C$, relations in (6.41) can be proved analogously to (6.29) and we omit the details. Consider (6.40). Split $\sigma_t^2(\theta) - \tilde{\sigma}_t^2(\theta) = U_{t,1}(\theta) + U_{t,2}(\theta)$, where

$$\begin{aligned} U_{t,1}(\theta) &:= \sum_{\ell=1}^{t-1} \gamma^\ell \{ (a + cY_{t-\ell}(d))^2 - (a + c\tilde{Y}_{t-\ell}(d))^2 \}, \\ U_{t,2}(\theta) &:= \sum_{\ell=t}^{\infty} \gamma^\ell \{ \omega^2 + (a + cY_{t-\ell}(d))^2 \}. \end{aligned} \quad (6.42)$$

Then $\sup_{\theta \in \Theta} |\partial^{\mathbf{j}} U_{t,i}(\theta)| \xrightarrow{p} 0, t \rightarrow \infty, i = 1, 2$ follows by using Assumption (B) and considering the bounds on the derivatives as in the proof of (6.29). For instance, let us prove (6.40) for $\partial^{\mathbf{j}} = \partial_d, |\mathbf{j}| = 1$. We have $|\partial_d U_{t,1}(\theta)| \leq C \sum_{\ell=1}^{t-1} \gamma^\ell \{ (1 + |\bar{Y}_{t-\ell}(d)|) |\partial_d(Y_{t-\ell}(d) - \tilde{Y}_{t-\ell}(d))| + |\partial_d Y_{t-\ell}(d)| |Y_{t-\ell}(d) - \tilde{Y}_{t-\ell}(d)| \}$. Hence, $\sup_{\theta \in \Theta} |\partial_d U_{t,1}(\theta)| \xrightarrow{p} 0$ follows from $0 \leq \gamma \leq \gamma_2 < 1$ and

$$\mathbb{E} \sup_{d \in [d_1, d_2]} (|Y_t(d) - \tilde{Y}_t(d)|^2 + |\partial_d(Y_t(d) - \tilde{Y}_t(d))|^2) \rightarrow 0 \quad \text{and} \quad (6.43)$$

$$\mathbb{E} \sup_{d \in [d_1, d_2]} (|Y_t(d)|^2 + |\tilde{Y}_t(d)|^2 + |\partial_d Y_t(d)|^2 + |\partial_d \tilde{Y}_t(d)|^2) \leq C. \quad (6.44)$$

The proof of (6.44) mimics that of (6.33) and therefore is omitted. To show (6.43), note $Y_t(d) - \tilde{Y}_t(d) = \sum_{j=t}^{\infty} j^{d-1} r_{t-j}$ and use a similar argument as in (6.33) to show that the l.h.s. of (6.44) does not exceed $C \sup_{d \in [d_1, d_2]} \sum_{i=0}^2 \mathbb{E} |\partial_d^i (Y_t(d) - \tilde{Y}_t(d))|^2 \leq C \sup_{d \in [d_1, d_2]} \sum_{j=t}^{\infty} j^{2(d-1)} (1 + \log^2 j) \rightarrow 0$ ($t \rightarrow \infty$). This proves (6.40) for $|j| = 1$ and $\partial^{\mathbf{j}} = \partial_d$. The remaining cases in (6.40) follow similarly and we omit the details. This proves (4.21) and completes the proof of Lemma 4.1. \square

Proof of Lemma 4.2. We have $|L(\theta_1) - L(\theta_2)| \leq \mathbb{E} |l_t(\theta_1) - l_t(\theta_2)| \leq C \mathbb{E} |\sigma_t^2(\theta_1) - \sigma_t^2(\theta_2)|$, where the last expectation can be easily shown to vanish as $|\theta_1 - \theta_2| \rightarrow 0$, $\theta_1, \theta_2 \in \Theta$. This proves the first statement of the lemma. To show the second statement of the lemma, write

$$L(\theta) - L(\theta_0) = \mathbb{E} \left[\frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - \log \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1 \right].$$

The function $f(x) := x - 1 - \log x > 0$ for $x > 0, x \neq 1$ and $f(x) = 0$ if and only if $x = 1$. Therefore $L(\theta) \geq L(\theta_0), \forall \theta \in \Theta$ while $L(\theta) = L(\theta_0)$ is equivalent to

$$\sigma_t^2(\theta) = \sigma_t^2(\theta_0) \quad (\mathbb{P}_{\theta_0} - \text{a.s.}) \quad (6.45)$$

Thus, it remains to show that (6.45) implies $\theta = \theta_0 = (\gamma_0, \omega_0, a_0, d_0, c_0)$. Consider the ‘projection’ $P_s \xi = \mathbb{E}[\xi | \mathcal{F}_s] - \mathbb{E}[\xi | \mathcal{F}_{s-1}]$ of r.v. $\xi, \mathbb{E}|\xi| < \infty$, where $\mathcal{F}_s = \sigma(\zeta_u, u \leq s)$ (see sec.2). (6.45) implies

$$0 = P_s(\sigma_t^2(\theta) - \sigma_t^2(\theta_0)) = P_s(Q_t^2(\theta) - Q_t^2(\theta_0)) + (\gamma - \gamma_0)P_s\sigma_{t-1}^2(\theta_0), \quad \forall s \leq t-1, \quad (6.46)$$

where $Q_t^2(\theta) = \omega^2 + (a + \sum_{u < t} b_{t-u}(\theta) r_u)^2$ is the same as in (6.35). We have

$$\begin{aligned} P_s Q_t^2(\theta) &= 2ab_{t-s}(\theta)r_s + 2b_{t-s}(\theta)r_s \sum_{u < s} b_{t-u}(\theta)r_u + \sum_{s \leq u < t} b_{t-u}^2(\theta)P_s r_u^2 \\ &= 2ab_{t-s}(\theta)\zeta_s \sigma_s(\theta_0) + 2b_{t-s}(\theta)\zeta_s \sigma_s(\theta_0) \sum_{u < s} b_{t-u}(\theta)r_u \\ &\quad + \sum_{s < u < t} b_{t-u}^2(\theta)P_s \sigma_u^2(\theta_0) + b_{t-s}^2(\theta)(\zeta_s^2 - 1)\sigma_s^2(\theta_0). \end{aligned} \quad (6.47)$$

Whence and from (6.46) for $s = t-1$ using $P_{t-1}\sigma_{t-1}^2(\theta_0) = 0$ we obtain

$$C_1(\theta, \theta_0)\zeta_{t-1}^2 + 2C_2(\theta, \theta_0)\zeta_{t-1} - C_1(\theta, \theta_0) = 0 \quad (6.48)$$

where

$$\begin{aligned} C_1(\theta, \theta_0) &:= (c^2 - c_0^2)\sigma_{t-1}(\theta_0), \\ C_2(\theta, \theta_0) &:= (ac - a_0c_0) + \sum_{u < t-1} (c^2(t-u)^{d-1} - c_0^2(t-u)^{d_0-1})r_u. \end{aligned}$$

Since $C_i(\theta, \theta_0), i = 1, 2$ are \mathcal{F}_{t-2} -measurable, (6.48) implies $C_1(\theta, \theta_0) = C_2(\theta, \theta_0) = 0$, particularly, $c = c_0$ since $\sigma_{t-1}(\theta_0) \geq \omega > 0$. Then $0 = C_2(\theta, \theta_0) = c_0(a - a_0) + c_0^2 \sum_{u < t-1} ((t-u)^{d-1} - (t-u)^{d_0-1})r_u$ and $\mathbb{E}r_u = 0$ lead to $a = a_0$ and next to $0 = \mathbb{E}(\sum_{u < t-1} ((t-u)^{d-1} - (t-u)^{d_0-1})r_u)^2 = \mathbb{E}r_0^2 \sum_{j \geq 2} (j^{d-1} - j^{d_0-1})^2 = 0$, or $d = d_0$. Consequently, $P_s(Q_t^2(\theta) - Q_t^2(\theta_0)) =$

0 for any $s \leq t-1$ and hence $\gamma = \gamma_0$ in view of (6.46). Finally, $\omega = \omega_0$ follows from $E\sigma_t^2(\theta) = E\sigma_t^2(\theta_0)$ and the fact that $\omega > 0, \omega_0 > 0$. This proves $\theta = \theta_0$ and the lemma, too. \square

Proof of Lema 4.3. From (3.17), it suffices to show that

$$\nabla \sigma_t^2(\theta) \lambda^T = 0 \quad (6.49)$$

for some $\theta \in \Theta$ and $\lambda \in \mathbb{R}^5, \lambda \neq 0$ leads to a contradiction. To the last end, we use a similar projection argument as in the proof of Lemma 4.2. First, note that $\sigma_t^2(\theta) = Q_t^2(\theta) + \gamma \sigma_{t-1}^2(\theta)$ implies

$$\nabla \sigma_t^2(\theta) = (0, \nabla_4 Q_t^2(\theta)) + \gamma \nabla \sigma_{t-1}^2(\theta) + (\nabla \gamma) \sigma_{t-1}^2(\theta),$$

where $\nabla_4 = (\partial/\partial\theta_2, \dots, \partial/\partial\theta_5)$. Hence and using the fact that (6.49) holds for any $t \in \mathbb{Z}$ by stationarity, from (6.49) we obtain

$$(\sigma_{t-1}^2(\theta), \nabla_4 Q_t^2(\theta)) \lambda^T = 0. \quad (6.50)$$

Thus,

$$(P_s \sigma_{t-1}^2(\theta), P_s \nabla_4^T Q_t^2(\theta)) \lambda = 0, \quad \forall s \leq t-1;$$

c.f. (6.46). For $s = t-1$ using $P_{t-1} \sigma_{t-1}^2(\theta) = 0$, $P_{t-1} \nabla_4 Q_t^2(\theta) = \nabla_4 P_{t-1} Q_t^2(\theta)$ by differentiating (6.47) similarly to (6.48) we obtain

$$D_1(\lambda) \zeta_{t-1}^2 + 2D_2(\lambda) \zeta_{t-1} - D_1(\lambda) = 0 \quad (6.51)$$

where $D_1(\lambda) := 2\lambda_5 \sigma_{t-1}(\theta)$ and

$$D_2(\lambda) := \lambda_3 c + \lambda_5 a + 2\lambda_5 c \sum_{u < t-1} (t-u)^{d-1} r_u + \lambda_4 c^2 \sum_{u < t-1} (t-u)^{d-2} \log(t-u) r_u,$$

$\lambda = (\lambda_1, \dots, \lambda_5)^T$. As in (6.48), $D_i(\lambda), i = 1, 2$ are \mathcal{F}_{t-2} -measurable, (6.51) implying $D_i(\lambda) = 0, i = 1, 2$. Hence, $\lambda_5 = 0$ and then $D_2(\lambda) = 0$ reduces to $\lambda_3 c + \lambda_4 c^2 \sum_{u < t-1} (t-u)^{d-2} \log(t-u) r_u = 0$. By taking expectation and using $c \neq 0$ we get $\lambda_3 = 0$ and then $\lambda_4 = 0$ since $E(\sum_{u < t-1} (t-u)^{d-2} \log(t-u) r_u)^2 \neq 0$. The above facts allow to rewrite (6.50) as $2\omega \lambda_2 + \lambda_1 \sigma_{t-1}^2(\theta) = 0$. Unless both λ_1, λ_2 vanish, the last equation means that either $\lambda_1 \neq 0$ and $\{\sigma_t^2(\theta)\}$ is a deterministic process which contradicts $c \neq 0$, or $\lambda_1 = 0, \lambda_2 \neq 0$ and $\omega = 0$, which contradicts $\omega \neq 0$. Lemma 4.3 is proved. \square

Proof of Lemma 4.4. Consider the first relation in (4.22). The pointwise convergence $L_n(\theta) \xrightarrow{a.s.} L(\theta)$ follows by ergodicity of $\{l_t(\theta)\}$ and the uniform convergence in (4.22) from $E \sup_{\theta \in \Theta} |\nabla l_t(\theta)| < \infty$, c.f. ([1], proof of Lemma 3), which in turn follows from of Lemma 4.1 (4.20) with $p = 1$. The proof of the second relation in (4.22) is immediate from Lemma 4.1 (4.21) with $p = 0, \epsilon = 1$. The proof of the statements (ii) and (iii) using Lemma 4.1 is similar and is omitted. \square

Proof of Theorem 4.1. (i) Follows from Lemmas 4.2 and 4.4 (i) using standard argument.

(ii) By Taylor's expansion,

$$0 = \nabla L_n(\hat{\theta}_n) = \nabla L_n(\theta_0) + \nabla^T \nabla L_n(\theta_n^*)(\hat{\theta}_n - \theta_0),$$

where $\theta_n^* \rightarrow_p \theta_0$ since $\hat{\theta}_n \rightarrow_p \theta_0$. Then $\nabla^T \nabla L_n(\theta_n^*) \rightarrow_p \nabla^T \nabla L(\theta_0)$ by Lemma 4.4 (4.24). Next, since $\{r_t^2/\sigma_t^2(\theta_0) - 1, \mathcal{F}_t, t \in \mathbb{Z}\}$ is a square-integrable and ergodic martingale difference sequence, the convergence $n^{1/2} \nabla L_n(\theta_0) \xrightarrow{law} N(0, A(\theta_0))$ follows by the martingale central limit theorem in ([2], Thm.23.1). Then (4.18) follows by Slutsky's theorem and (3.16). \square

Proof of Theorem 4.2. Part (i) follows from Lemmas 4.2 and 4.4 (i) as in the case of Theorem 4.1 (i). To prove part (ii), by Taylor's expansion

$$0 = \nabla \tilde{L}_n^{(\beta)}(\tilde{\theta}_n^{(\beta)}) = \nabla \tilde{L}_n^{(\beta)}(\theta_0) + \nabla^T \nabla \tilde{L}_n^{(\beta)}(\tilde{\theta}_n^*)(\tilde{\theta}_n^{(\beta)} - \theta_0),$$

where $\tilde{\theta}_n^* \rightarrow_p \theta_0$ since $\tilde{\theta}_n^{(\beta)} \rightarrow_p \theta_0$. Then $\nabla^T \nabla \tilde{L}_n^{(\beta)}(\tilde{\theta}_n^*) \rightarrow_p \nabla^T \nabla L(\theta_0)$ by Lemma 4.4 (4.24)-(4.25). From the proof of Theorem 4.1 (ii) we have that $n^{\beta/2} \nabla L_n^{(\beta)}(\theta_0) \xrightarrow{law} N(0, A(\theta_0))$, where $L_n^{(\beta)}(\theta) := \frac{1}{[n^\beta]} \sum_{t=n-[n^\beta]+1}^n l_t(\theta)$. Hence, the central limit theorem in (4.19) follows from

$$I_n(\beta) := E|\nabla \tilde{L}_n^{(\beta)}(\theta_0) - \nabla L_n^{(\beta)}(\theta_0)| = o(n^{-\beta/2}). \quad (6.52)$$

We have $I_n(\beta) \leq \sup_{n-[n^\beta] \leq t \leq n} E|\nabla l_t(\theta_0) - \nabla \tilde{l}_t(\theta_0)|$ and (6.52) follows from

$$E|\nabla l_t(\theta_0) - \nabla \tilde{l}_t(\theta_0)| = o(t^{-\beta/2}), \quad t \rightarrow \infty. \quad (6.53)$$

Write $\|\xi\|_p := E^{1/p}|\xi|^p$ for L^p -norm of r.v. ξ . Using $|\nabla(l_t(\theta_0) - \tilde{l}_t(\theta_0))| \leq r_t^2 |\nabla(\sigma_t^{-2}(\theta_0) - \tilde{\sigma}_t^{-2}(\theta_0))| + |\nabla(\log \sigma_t^2(\theta_0) - \log \tilde{\sigma}_t^2(\theta_0))|$ and assumption $E|r_t|^5 < \infty$, relation (6.53) follows from

$$\begin{aligned} \|\sigma_t^{-4} \partial_i \sigma_t^2 - \tilde{\sigma}_t^{-4} \partial_i \tilde{\sigma}_t^2\|_{5/3} &= O(t^{d_0-1/2} \log t) \quad \text{and} \\ \|\sigma_t^{-2} \partial_i \sigma_t^2 - \tilde{\sigma}_t^{-2} \partial_i \tilde{\sigma}_t^2\|_1 &= O(t^{d_0-1/2} \log t), \quad i = 1, \dots, 5, \end{aligned} \quad (6.54)$$

where $\sigma_t^2 := \sigma_t^2(\theta_0)$, $\tilde{\sigma}_t^2 := \tilde{\sigma}_t^2(\theta_0)$, $\partial_i \sigma_t^2 := \partial_i \sigma_t^2(\theta_0)$, $\partial_i \tilde{\sigma}_t^2 := \partial_i \tilde{\sigma}_t^2(\theta_0)$. Below, we prove the first relation (6.54) only, the proof of the second one being similar. We have $\sigma_t^{-4} \partial_i \sigma_t^2 - \tilde{\sigma}_t^{-4} \partial_i \tilde{\sigma}_t^2 = \sigma_t^{-4} \tilde{\sigma}_t^{-4} (\tilde{\sigma}_t^2 + \sigma_t^2)(\tilde{\sigma}_t^2 - \sigma_t^2) \partial_i \sigma_t^2 + \tilde{\sigma}_t^{-4} (\partial_i \sigma_t^2 - \partial_i \tilde{\sigma}_t^2)$. Then using $\sigma_t^2 \geq \omega_1^2/(1-\gamma_2) > 0$, $\tilde{\sigma}_t^2 \geq \omega_1^2/(1-\gamma_2) > 0$, relation the first relation in (6.54) follows from

$$\|(\sigma_t^2 - \tilde{\sigma}_t^2)(\partial_i \sigma_t^2 / \sigma_t)\|_{5/3} = O(t^{d_0-1/2}) \quad \text{and} \quad (6.55)$$

$$\|\partial_i \sigma_t^2 - \partial_i \tilde{\sigma}_t^2\|_{5/3} = O(t^{d_0-1/2} \log t), \quad i = 1, \dots, 5. \quad (6.56)$$

Consider (6.55). By Hölder's inequality, $\|(\sigma_t^2 - \tilde{\sigma}_t^2)(\partial_i \sigma_t^2 / \sigma_t)\|_{5/3} \leq \|\sigma_t^2 - \tilde{\sigma}_t^2\|_{5/2} \|\partial_i \sigma_t^2 / \sigma_t\|_5$, where $\|\partial_i \sigma_t^2 / \sigma_t\|_5 < C$ according to (6.29). Hence, (6.55) follows from

$$\|\sigma_t^2 - \tilde{\sigma}_t^2\|_{5/2} = O(t^{d_0-1/2}). \quad (6.57)$$

To show (6.57), similarly as in the proof of (6.40) split $\sigma_t^2 - \tilde{\sigma}_t^2 = U_{t,1} + U_{t,2}$, where $U_{t,i} := U_{t,i}(\theta_0)$, $i = 1, 2$ are defined in (6.42), i.e., $U_{t,1} = \sum_{\ell=1}^{t-1} \gamma_0^\ell \{(a_0 + c_0 Y_{t-\ell})^2 - (a_0 + c_0 \tilde{Y}_{t-\ell})^2\}$, $U_{t,2} = \sum_{\ell=t}^\infty \gamma_0^\ell \{\omega_0^2 + (a_0 + c_0 Y_{t-\ell})^2\}$ and $Y_t := Y_t(d_0)$, $\tilde{Y}_t := \tilde{Y}_t(d_0)$. We have $|U_{t,1}| \leq C \sum_{\ell=1}^{t-1} \gamma_0^\ell |Y_{t-\ell} - \tilde{Y}_{t-\ell}|(1 + |Y_{t-\ell}| + |\tilde{Y}_{t-\ell}|)$, $|U_{t,2}| \leq C \sum_{\ell=t}^\infty \gamma_0^\ell (1 + |Y_{t-\ell}|^2)$ and

hence

$$\begin{aligned}\|\sigma_t^2 - \tilde{\sigma}_t^2\|_{5/2} &\leq C \left\{ \sum_{\ell=1}^{t-1} \gamma_0^\ell \|Y_{t-\ell} - \tilde{Y}_{t-\ell}\|_5 (1 + |Y_{t-\ell}| + |\tilde{Y}_{t-\ell}|) \|_{5/2} + \sum_{\ell=t}^{\infty} \gamma_0^\ell (1 + \|Y_{t-\ell}\|_5) \right\} \\ &\leq C \left\{ \sum_{\ell=1}^{t-1} \gamma_0^\ell \|Y_{t-\ell} - \tilde{Y}_{t-\ell}\|_5 + \sum_{\ell=t}^{\infty} \gamma_0^\ell \right\},\end{aligned}\tag{6.58}$$

where we used the fact that $\|Y_t\|_5 < C$, $\|\tilde{Y}_t\|_5 < C$ by $\|r_t\|_5 < C$ and Rosenthal's inequality in (2.5). In a similar way from (2.5) it follows that

$$\|Y_{t-\ell} - \tilde{Y}_{t-\ell}\|_5 \leq C \left\{ \sum_{j>t-\ell} j^{2(d_0-1)} \right\}^{1/2} \leq C(t-\ell)^{d_0-1/2}.\tag{6.59}$$

Substituting (6.59) into (6.58) we obtain

$$\|\sigma_t^2 - \tilde{\sigma}_t^2\|_{5/2} \leq C \left\{ \sum_{\ell=1}^{t-1} \gamma_0^\ell (t-\ell)^{d_0-1/2} + \sum_{\ell=t}^{\infty} \gamma_0^\ell \right\} = O(t^{d_0-1/2}),$$

proving (6.57).

It remains to show (6.56). Similarly as above, $\partial_i \sigma_t^2 - \partial_i \tilde{\sigma}_t^2 = \partial_i U_{t,1} + \partial_i U_{t,2}$, where $\partial_i U_{t,j} := \partial_i U_{t,j}(\theta_0)$, $j = 1, 2$. Then (6.56) follows from

$$\|\partial_i U_{t,1}\|_{5/3} = O(t^{d_0-1/2} \log t) \quad \text{and} \quad \|\partial_i U_{t,2}\|_{5/3} = o(t^{d_0-1/2}), \quad i = 1, \dots, 5.\tag{6.60}$$

For $i = 1$, the proof of (6.60) is similar to (6.58). Consider (6.60) for $2 \leq i \leq 5$. Denote $V_t(\theta) := 2a + c(Y_t(d) + \tilde{Y}_t(d))$, $V_t := V_t(\theta_0)$, $\partial_i V_t := \partial_i V_t(\theta_0)$, then

$$\|\partial_i U_{t,1}\|_{5/3} \leq C \sum_{\ell=1}^{t-1} \gamma_0^\ell \left\{ \|\partial_i(Y_{t-\ell} - \tilde{Y}_{t-\ell})\|_5 \|V_t\|_5 + \|Y_{t-\ell} - \tilde{Y}_{t-\ell}\|_5 \|\partial_i V_t\|_5 \right\},$$

where $\partial_i(Y_{t-\ell} - \tilde{Y}_{t-\ell}) = 0$, $\partial_i \neq \partial_d$ and

$$\begin{aligned}\|\partial_d(Y_t - \tilde{Y}_t)\|_5 &= \left\| \sum_{j>t} j^{d_0-1} (\log j) r_{t-j} \right\|_5 \\ &\leq C \left\{ \sum_{j>t} j^{2(d_0-1)} \log^2 j \right\}^{1/2} = O(t^{d_0-1/2} \log t)\end{aligned}$$

similarly as in (6.59) above. Hence, the first relation in (6.60) follows from (6.59) and $\|\partial_i V_t\|_5 \leq C(1 + \|\partial_d Y_{t-\ell}\|_5 + \|\partial_d \tilde{Y}_{t-\ell}\|_5) \leq C < \infty$ as in the proof of (6.56), and the proof of the second relation in (6.60) is analogous. This proves (6.53) and completes the proof of Theorem 4.2. \square

Acknowledgement

This research was supported in part by grant MIP-063/2013 from the Research Council of Lithuania.

References

- [1] Beran, J., Schützner, M., 2009. On approximate pseudo-maximum likelihood estimation for LARCH-processes. *Bernoulli*, 15, 1057–1081.
- [2] Billingsley, P., 1968. *Convergence of Probability Measures*. New York: Wiley.
- [3] Burkholder, D.L., 1973. Distribution functions inequalities for martingales. *Ann. Probab.*, 1, 19–42.
- [4] Doukhan, P., Grublyté, I., Surgailis, D., 2014. A nonlinear model for long memory conditional heteroscedasticity. Preprint. Available at arXiv:1502.00095 [math.ST]
- [5] Engle, R.F., 1990. Stock volatility and the crash of '87. *Discussion. Rev. Financial Studies* 3, 103–106.
- [6] Francq, C., Zakoian, J.-M., 2010. Inconsistency of the MLE and inference based on weighted LS for LARCH models. *J. Econometrics* 159, 151–165.
- [7] Giraitis, L., Robinson, P.M., Surgailis, D., 2000. A model for long memory conditional heteroskedasticity. *Ann. Appl. Probab.*, 10, 1002–1024.
- [8] Grublyté, I., Škarnulis, A., 2015. A generalized nonlinear model for long memory conditional heteroscedasticity. Preprint. Available at arXiv:1509.01708 [math.ST]
- [9] Levine, M., Torres, S., Viens, F., 2009. Estimation for the long-memory parameter in LARCH models, and fractional Brownian motion. *Stat. Inf. Stoch. Process.* 12, 221–250.
- [10] Osękowski, A., 2012. A note on Burkholder-Rosenthal inequality. *Bull. Polish Acad. Sci. Mathematics* 60, 177–185.
- [11] Robinson, P.M., 1991. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. Econometrics*, 47, 67–84.
- [12] Rosenthal, H.P., 1970. On the subspaces of L^p ($p > 2$) spanned by the sequences of independent random variables. *Israel J. Math.*, 8, 273–303.
- [13] Sentana, E., 1995. Quadratic ARCH models. *Rev. Econom. Stud.* 3, 77–102.
- [14] Truquet, L., 2014. On a family of contrasts for parametric inference in degenerate ARCH models. *Econometric Th.*, 30, 1165–1206.

Table 1: Sample RMSE of QML estimates of $\theta_0 = (\gamma_0, \omega_0, a_0, c_0, d_0)$ of the GQARCH process in (1.3) for $\gamma_0 = 0.7, a_0 = -0.2, c_0 = 0.2$ and different values of ω_0, d_0 . The number of replications is 100

n	d_0	$\omega_0=0.1$				
		$\hat{\gamma}_n$	$\hat{\omega}_n$	\hat{a}_n	\hat{d}_n	\hat{c}_n
1000	0.1	0.091	0.057	0.035	0.103	0.035
	0.2	0.083	0.047	0.045	0.109	0.031
	0.3	0.071	0.045	0.047	0.094	0.043
	0.4	0.073	0.029	0.054	0.097	0.036
5000	0.1	0.031	0.021	0.012	0.047	0.015
	0.2	0.030	0.015	0.015	0.041	0.014
	0.3	0.028	0.011	0.025	0.042	0.013
	0.4	0.031	0.014	0.053	0.059	0.018
n	d_0	$\omega_0=0.01$				
		$\hat{\gamma}_n$	$\hat{\omega}_n$	\hat{a}_n	\hat{d}_n	\hat{c}_n
1000	0.1	0.070	0.049	0.030	0.103	0.029
	0.2	0.061	0.043	0.035	0.089	0.024
	0.3	0.066	0.040	0.045	0.106	0.044
	0.4	0.055	0.042	0.056	0.105	0.038
5000	0.1	0.025	0.032	0.011	0.035	0.013
	0.2	0.022	0.028	0.013	0.032	0.013
	0.3	0.025	0.028	0.025	0.046	0.016
	0.4	0.031	0.031	0.046	0.096	0.034
n	d_0	$\omega_0=0.001$				
		$\hat{\gamma}_n$	$\hat{\omega}_n$	\hat{a}_n	\hat{d}_n	\hat{c}_n
1000	0.1	0.086	0.058	0.026	0.095	0.037
	0.2	0.056	0.043	0.027	0.084	0.031
	0.3	0.053	0.039	0.046	0.080	0.029
	0.4	0.055	0.047	0.060	0.122	0.041
5000	0.1	0.022	0.033	0.009	0.031	0.012
	0.2	0.020	0.030	0.012	0.028	0.012
	0.3	0.022	0.032	0.024	0.038	0.014
	0.4	0.032	0.037	0.046	0.098	0.031